Cellmark Forensic Services

NIST DNA Analyst Webinar Series

Validation and use of likeLTD

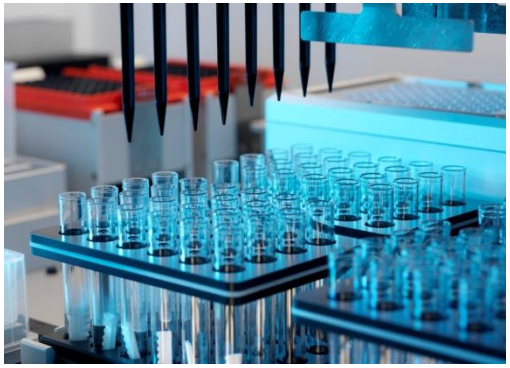
Matthew Greenhalgh
Director of Forensic Science

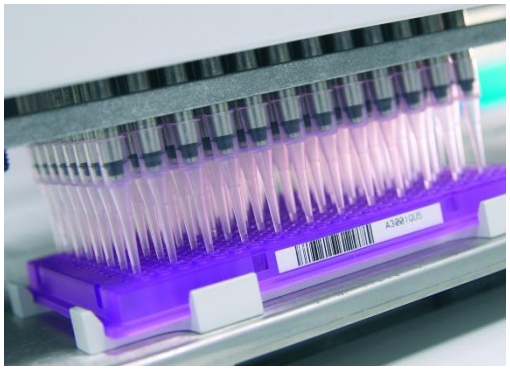
18 September 2014

Cellmark Forensic Services

Twenty Five Years of DNA Analysis

- Established ICI (Astrazeneca) in 1987
- Specialist Forensic DNA analysis
- One of Europe's largest paternity testers
- Contracted to >80% of UK police forces
- Approx 475 UK employees
- A LabCorp company since 2011



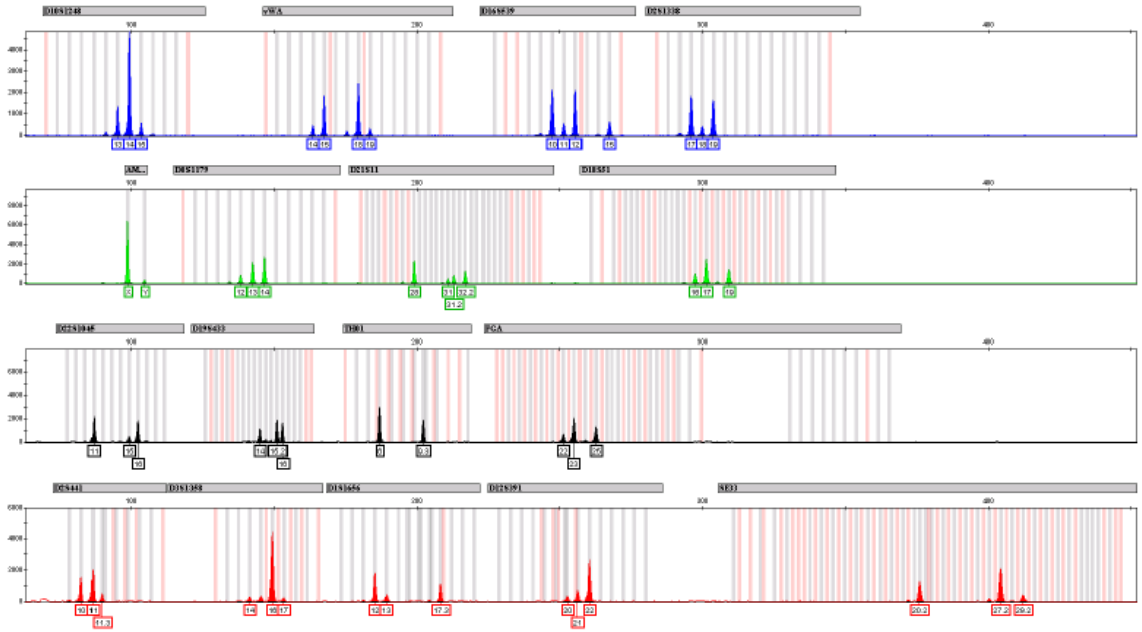


IDENTIFICATION INTERPRETATION INNOVATION

Cellmark Forensic Services

Background to the problem

- The increasing sensitivity of DNA analysis methods and their use in a wider range of case types has resulted in more mixtures:

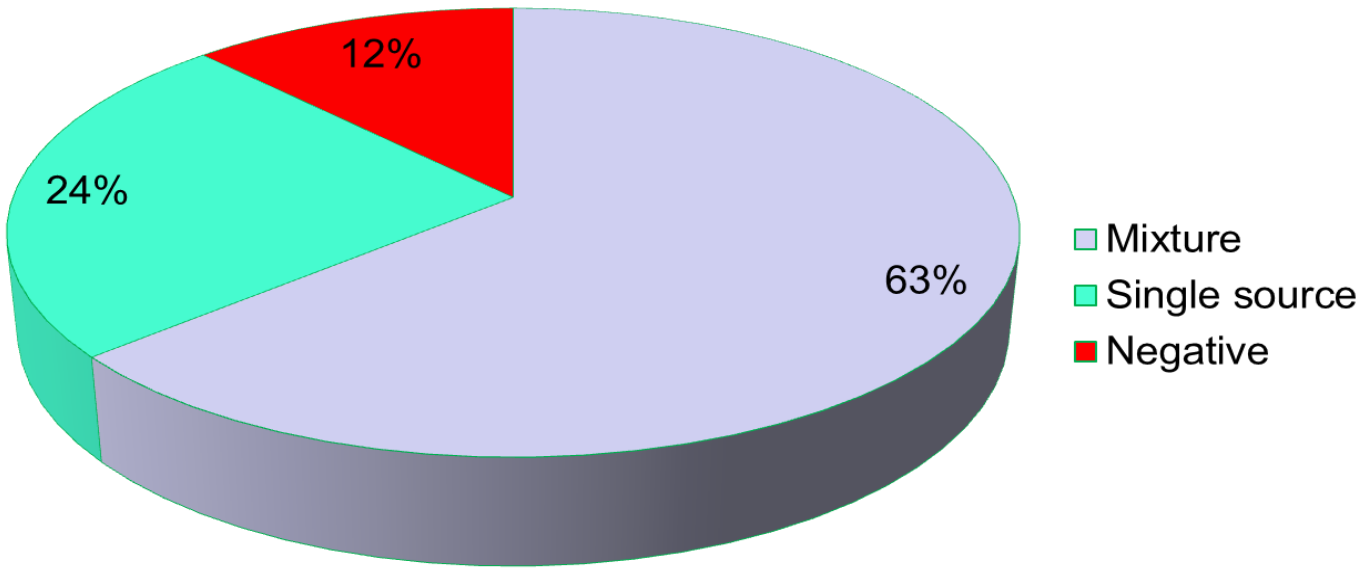


Cellmark Forensic Services

Interpretation procedures prior to likeLTD

- Match probability for single source profiles (or major/minor)
- LR method for 2 person mixtures where both individuals are fully represented
- “Dlugosz” expert opinions without statistical evaluation of the match in a limited number of cases
  - Court of Appeal Ruling R v Dlugosz, R v Pickering and R v MDS ([2013] EWCA Crim 2)

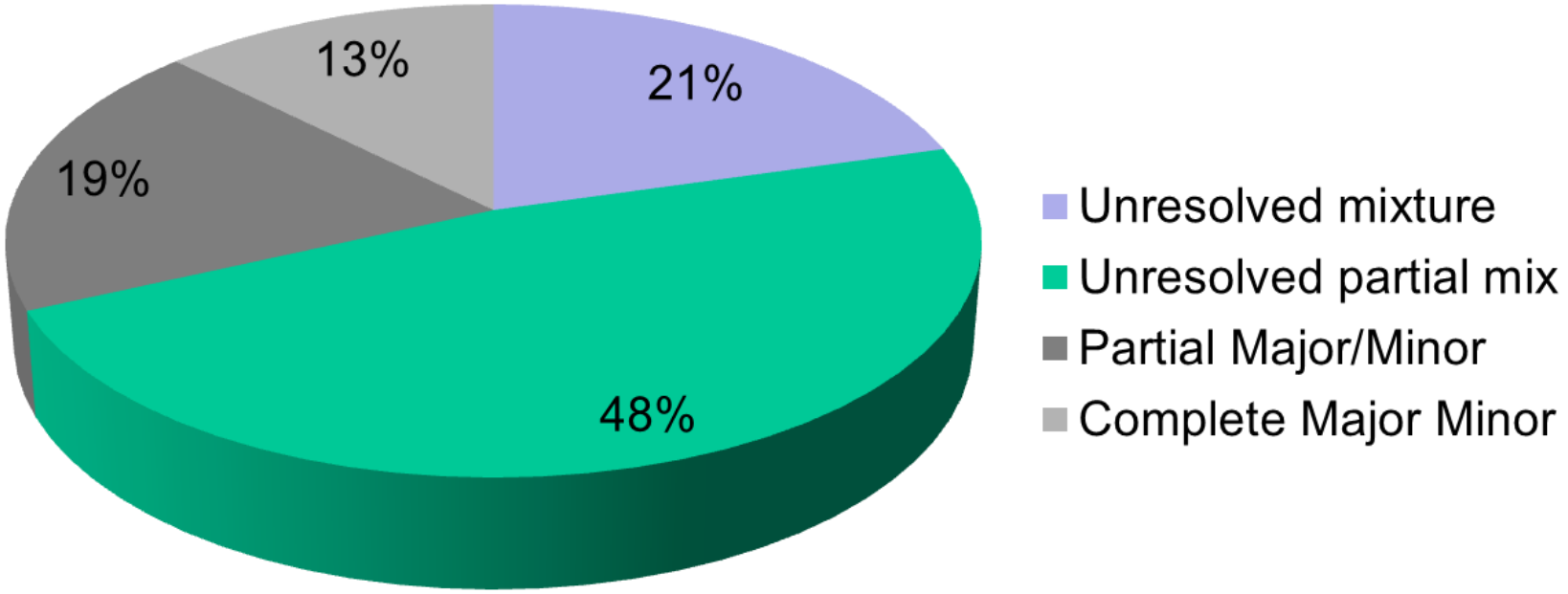
Distribution of DNA results - volume crime



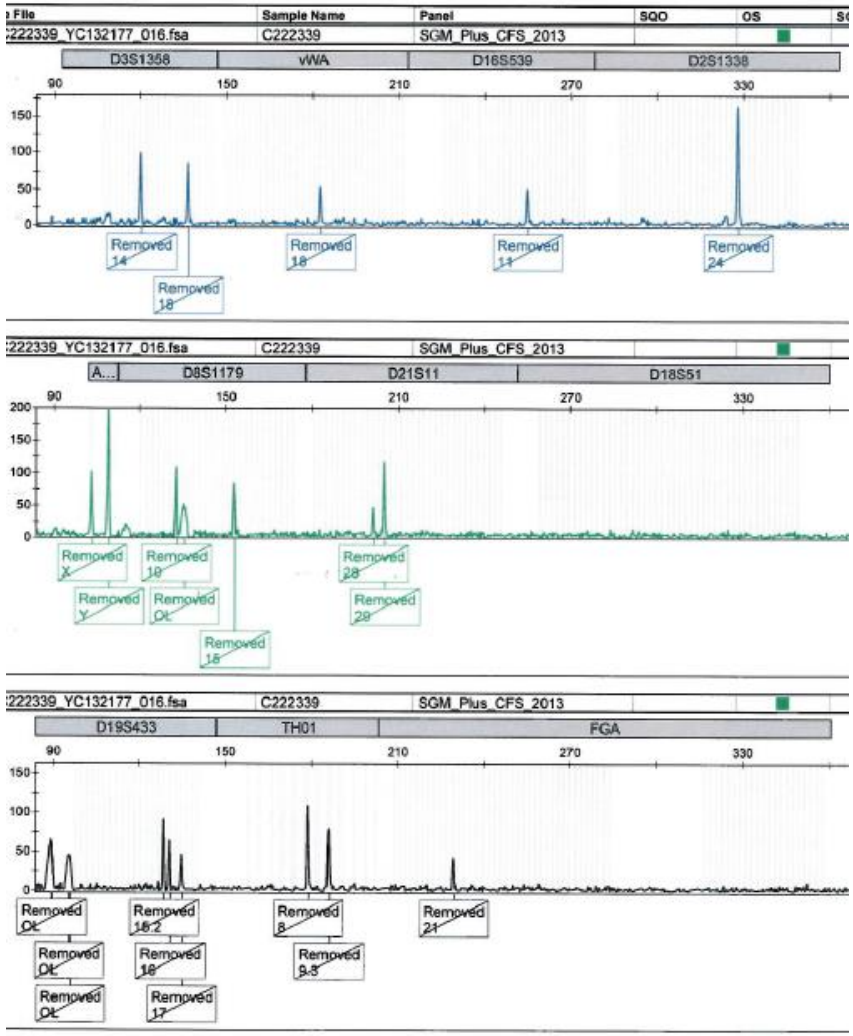
Based on 9165 samples (3 months data)

Cellmark Forensic Services

Breakdown of mixture results

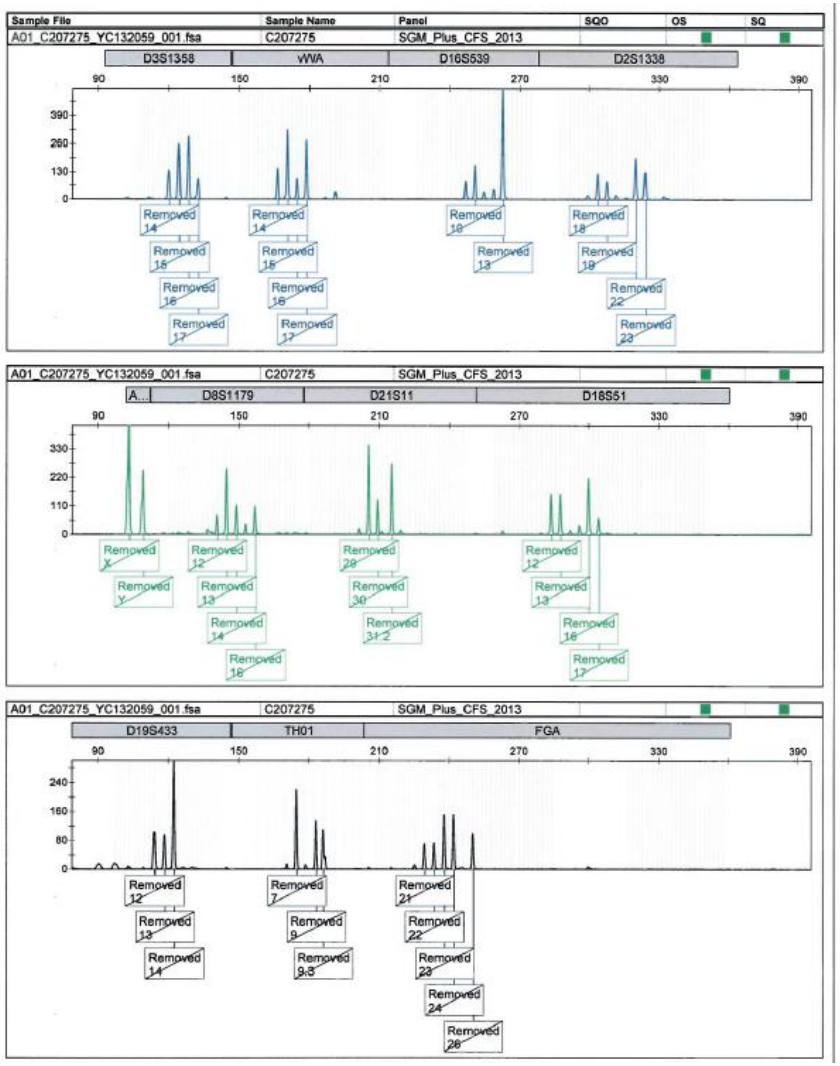


Profiles suitable for analysis



- Low level mixture
- 2 contributors?
- Suspect + unknown
- Allele dropout
- 3 replicates

Profiles suitable for analysis (*continued*)



- 3 contributors?
- Suspect + victim + unknown
- No apparent dropout
- 2 replicates

Cellmark Forensic Services

DNA Resolve – Cellmark Forensic Services

- Uses likeLTD software written by Prof. Balding (UCL London) combined with a Cellmark user-interface
- Models allele drop-out using Tvedebrink statistical model
- Option to include allele drop-in
- Can deal with a maximum of 2 unprofiled contributors in a mixture
- Multiple replicates can be analysed

likeLTD input files (.csv)

- Crime stain file

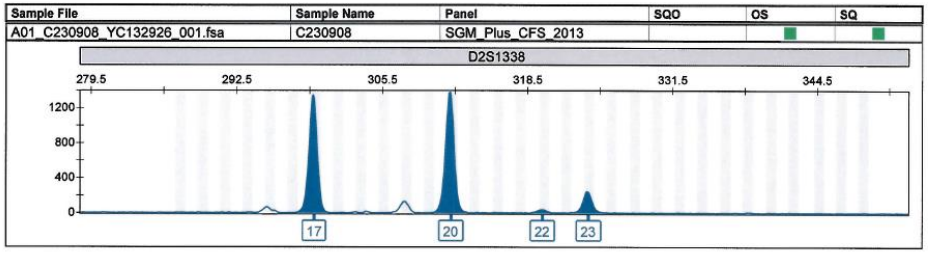
Stain	Profiling system	Plate/Run		D3	vWA	D16	D2	D8
CSP	SGM+	F2	Allelic	14,15,16,17,18	14,15,16,17,18	10,11,12,13	17,18,24	10,13,14,15
CSP	SGM+		Uncertain	13			16	
CSP	SGM+	G2	Allelic	14,16	15,16,17,18	10,11,12,13	17,18,25	10,12,13,14,15
CSP	SGM+		Uncertain	15		9	16,19	
CSP	SGM+	H2	Allelic	14,15,16,17	14,16	10,11,12,13	17,18	10,13,14,15
CSP	SGM+		Uncertain		15		25	
CSP	SGM+	B3	Allelic	14,15,16,17,18	14,15,16,17,18,19	9,10,11,13	17,18,20,25	10,12,13,14,15
CSP	SGM+		Uncertain	13		12	24	

- Reference File

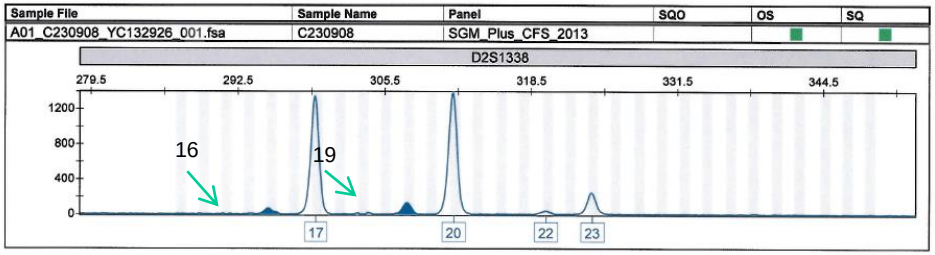
Individual	known/queried	D3	vWA	D16	D2	D8	D21
Suspect	queried	14,16	14,17	11,13	18,24	12,14	28,31.2
Victim	known	15	16,20	9,11	23	12,13	28,33.2

Allelic and Uncertain peaks

- Allelic Peaks



- Uncertain Peaks
  - Sub-threshold
  - Possible Stutter



Cellmark Forensic Services

Allele report

Data provided by forensic scientist

Crime scene profiles (CSP)

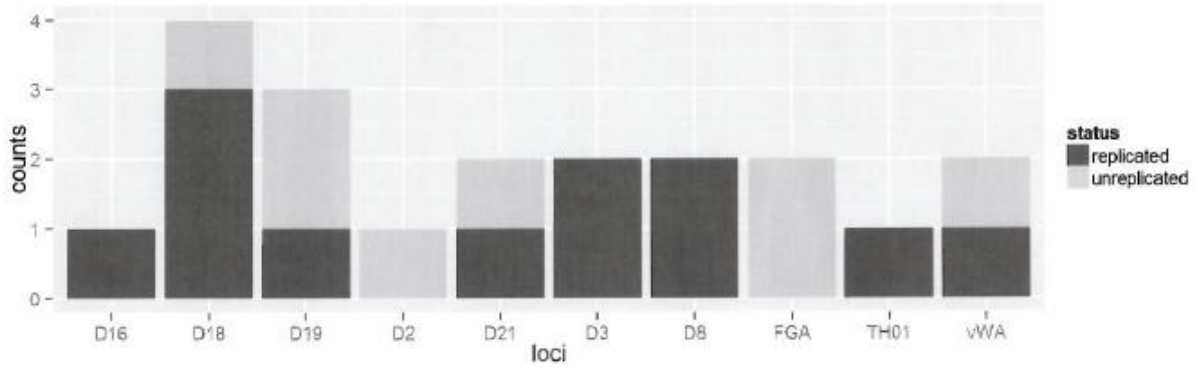
run	status	D3	vWA	D16	D2	D8	D21	D18	D19	TH01	FGA
1	certain	14,15,16,17,18	14,15,16,17,18	10,11,12,13	17,18,24	10,13,14,15	27,28,32.2	12,13,14,17	13,14	6,7,9,9.3	21,23,25
-	uncertain	13			16			11,16	12		20,26
2	certain	14,16	15,16,17,18	10,11,12,13	17,18,25	10,12,13,14,15	28,30,31,31.2,32.2	16,17	13,14,15	6,7,9	19,21,22,23
-	uncertain	15		9	16,19		27				
3	certain	14,15,16,17	14,16	10,11,12,13	17,18	10,13,14,15	28,32.2	12,14,15,17,18	13,14,14.2	6,7,9,9.3	21,22,23
-	uncertain		15		25			16			20,25,26
4	certain	14,15,16,17,18	14,15,16,17,18,19	9,10,11,13	17,18,20,25	10,12,13,14,15	27,28,32.2	13,16,17	13,14,15,15.2	6,7,9,9.3	20,21,22,23,25
-	uncertain	13		12	24						24

Reference profiles

profile	D3	vWA	D16	D2	D8	D21	D18	D19	TH01	FGA
Q	**14,16**	**15,16**	**10,13**	**17,18**	**10,15**	**28,32.2**	**17**	**14**	**6,7**	**21,23**
K	**15**	**14,18**	**11,*9***	**25,*24***	**14**	*30,31.2*	**13,*15***	**14,15**	**7,9**	**22,25**
Other	**17,18**	**17,*19***	**12**	*20*	**13,12**	**27,*31***	**12,14,16,*18***	**13,*14.2,15.2***	**9.3**	*19,20*

Alleles that are **replicated**, *unreplicated* or absent in the crime scene profile, using the certain designations only.

Allele report (*continued*)



The number of 'certain' alleles that cannot be attributed to the known profile(s).

Unusual alleles

source	locus	allele	EA1.freq	EA3.freq	EA4.freq	error
Crime scene uncertain	D3	13	5	7	0	-
Crime scene uncertain	D18	11	10	1	6	-

Alleles are automatically checked against the database. An error will be reported if an allele is absent from the database, or present more than once, or if a locus is absent.

Approximate representation

Contributor	Rep 1	Rep 2	Rep 3	Rep 4	Total
Q	100	100	94	100	99
K	61	61	50	78	62

Cellmark Forensic Services

Allele report (*continued*)

The fraction of an individual's alleles (as a percentage) that have been designated as 'certain' alleles in each replicate. This estimate is not used by likeLTD, and is intended to assist informal assessments of possible known contributors to the CSP. A more formal approach is to do a likeLTD run to compute the likelihood ratio (LR) for that individual contributor.

Suggested parameter values

nUnknowns	doDropin	Recommendation
2	FALSE	recommended

Recommended values for 'nUnknowns', choose from 0,1 or 2 (likeLTD automatically adds and additional unknown X to the defence hypothesis in place of the queried profile Q).
Recommended values for 'doDropin', choose from 'TRUE' or 'FALSE'.
All the attributable alleles must either come from an unknown or dropin.

System information

Type	Details
Date report generated:	Fri Aug 01 09:07:22 2014
Package	likeLTD
Title	Tools to determine DNA profile evidence.
Description	Tools to determine DNA profile Weight of Evidence. For further information see the likeLTD guide at the URL provided, or the paper under citation.
Depends	R (≥ 2.10), DEoptim, ggplot2, gtools, rtf
Suggests	svUnit, scales
Imports	gdata, tools, tcltk
Version	5.4.0
Date	2013-03-15
Author	David Balding, Adrian Timpson, Christopher Steele, Mayeul d'Avezac, James Hetherington.
Maintainer	Christopher Steele <c.steele.11@ucl.ac.uk>
License	GPL-3
URL	https://sites.google.com/site/baldingstatisticalgenetics/

Evaluation report

CFS-920055-14-ver5.4-Evaluation-Report

CFS-920055-14-ver5.4

Prosecution Hypothesis: Q (Q) + K + 2U

Defence Hypothesis: Unknown (X) + K + 2U

Cellmark Forensic Services

Evaluation report (*continued*)

The fraction of an individual's alleles (as a percentage) that have been designated as 'certain' alleles in each replicate. This estimate is not used by likeLTD, and is intended to assist informal assessments of possible known contributors to the CSP. A more formal approach is to do a likeLTD run to compute the likelihood ratio (LR) for that individual contributor.

Likelihoods at each locus

Likelihood	D3	vWA	D16	D2	D8	D21	D18	D19	TH01	FGA
Prosecution.log10	-2.443	-4.427	-1.216	-3.158	-1.751	-6.926	-9.377	-6.054	-1.503	-4.713
Defence.log10	-3.432	-5.615	-2.404	-4.444	-2.993	-8.446	-9.455	-5.790	-2.039	-5.821
Ratio.log10	0.989	1.187	1.188	1.286	1.241	1.520	0.078	-0.265	0.537	1.108
Ratio	9.739	15.397	15.407	19.324	17.432	33.120	1.196	0.543	3.441	12.827

Overall Likelihood

calculation	estimate
Prosecution.log10	-41.568
Defence.log10	-50.437
Ratio.log10	8.869
Ratio	739308745

Theoretical maximum LR

calculation	estimate
likelihood ratio	10915470994030
Log10 likelihood ratio	13.038

Dropout and degradation parameter estimates

Cellmark Forensic Services

Input required using R interface

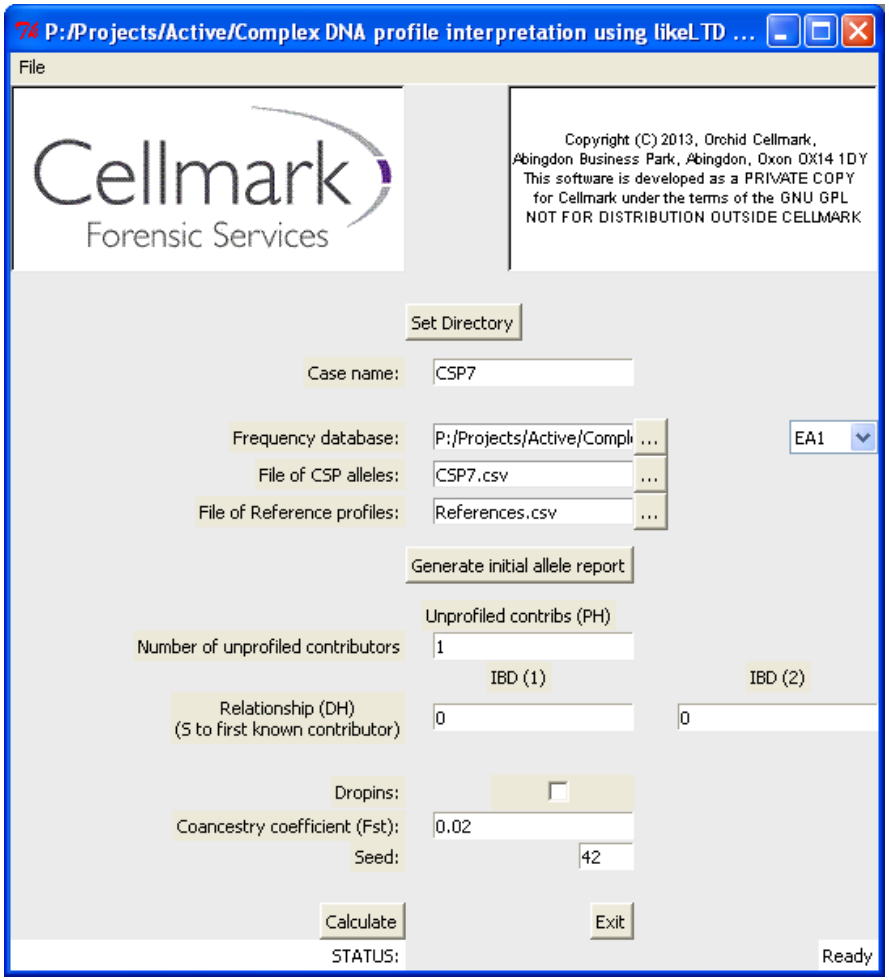
```
> require(likeLTD)
Loading required package: likeLTD
Loading required package: DEoptim

DEoptim package
Differential Evolution algorithm in R
Authors: D. Ardia, K. Mullen, B. Peterson and J. Ulrich

Loading required package: ggplot2
Loading required package: gtools
Loading required package: rtf
> setwd("P:/Projects/Active/Complex DNA profile interpretation using likeLTD software/
> datapath <- "."
> admin <- pack.admin.input(
+   cspFile = "CSP1.csv",
+   refFile = "References.csv",
+   databaseFile = "Cellmark_NDNAD-allele-freqs-DNA17_v2_(SGM PLUS ONLY)-wbp1.txt",
+   caseName = "CSP1 batch test"
+   )
> allele.report(admin)
>
> args <- list(
+   nUnknowns = 0,
+   doDropin = FALSE,
+   ethnic = "EA1",
+   adj = 1,
+   fst = 0.02,
+   relatedness = c(0,0)
+   )
>
> hypP <- do.call(prosecution.hypothesis, append(admin,args))
> hypD <- do.call(defence.hypothesis, append(admin,args))
>
> paramsP <- optimisation.params(hypP)
> paramsD <- optimisation.params(hypD)
>
> results <- evaluate(paramsP, paramsD, progBar = FALSE, interim = FALSE)
```

Cellmark
Forensic Services

Cellmark user interface



Cellmark Forensic Services

likeLTD validation

- Establishing that GUI did not affect results
- Repeat tests published by Balding
- Additional testing
- Validated under ISO 9001 certification
- Planning to add to ISO 17025 scope

Cellmark Forensic Services

Tests from likeLTD guide

CSP1 Full profile match to reference Q, single contributor

CSP2 The two replicates of CSP2 differ from reference Q due to 1 drop in and 2 dropouts

CSP3 One further drop in and two more dropouts have been introduced

CSP4 Two contributors: All the alleles of both contributors present in both replicates with no drop in or dropout

CSP5 Introduces random 50% dropout for the alleles of unknown 1 (U1) not shared with Q

Cellmark Forensic Services

Tests from likeLTD guide (*continued*)

CSP6 The opposite situation is considered where there is 50% dropout of the alleles of Q not shared with U1

CSP7 In addition to the 50% dropout for the alleles of Q, 50% of the alleles of U1 generate stutter peaks that are classified as uncertain

CSP8 Random 50% dropout affects both the alleles of Q and U1

Cellmark Forensic Services

Additional validation tests

- Effects on the LR value of increasing levels of Fst for Caucasian, Afro Caribbean and Asian frequency databases + comparison against the reciprocal of the match probability calculated using Cellmark's in house software
- Effect of analysis assuming the defence scenario that the donor of the DNA is the suspect's brother
- Effect of using an incorrect number of contributors

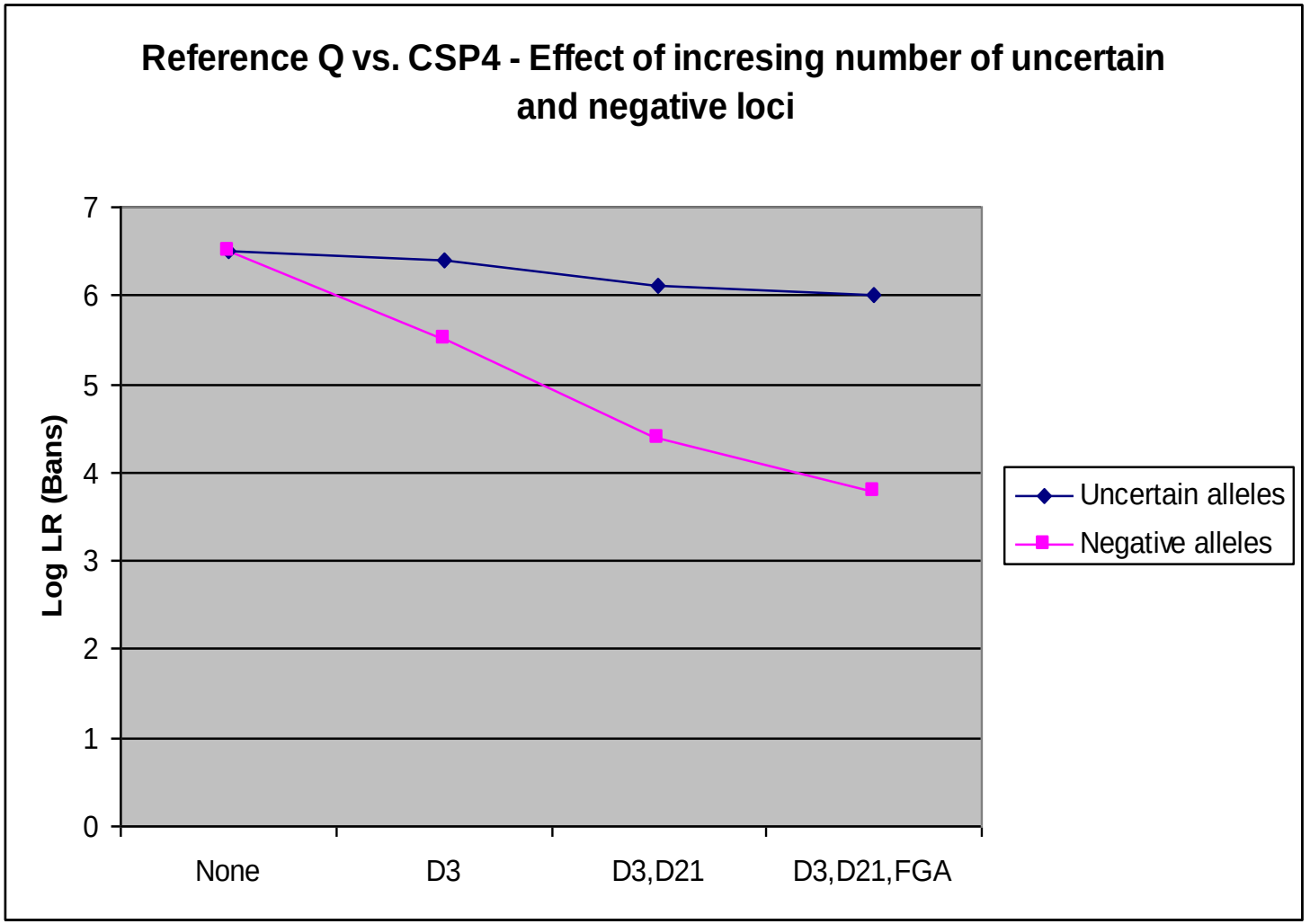
Cellmark Forensic Services

Additional validation tests (*continued*)

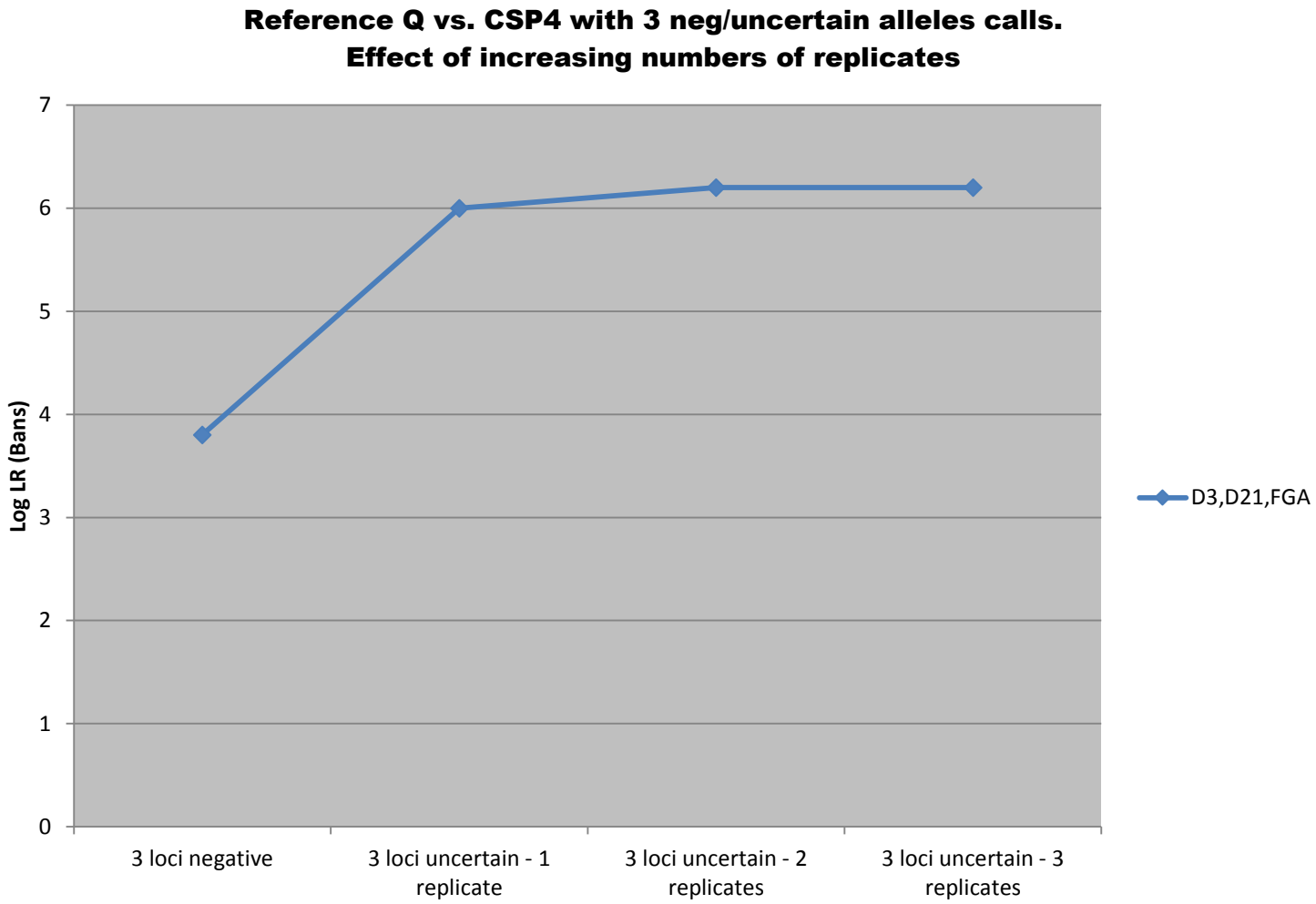
- Comparison of likeLTD against Mixture Analyser software
- Comparison of likeLTD against another probabilistic software package (STRmix)
- Effects of using the "uncertain" option for allele calls and varying the number of replicates
- Random reference profiles compared against crime scene stain
  - Single contributor
  - Two contributors

Cellmark Forensic Services

Effect of increasing the number of uncertain and negative loci

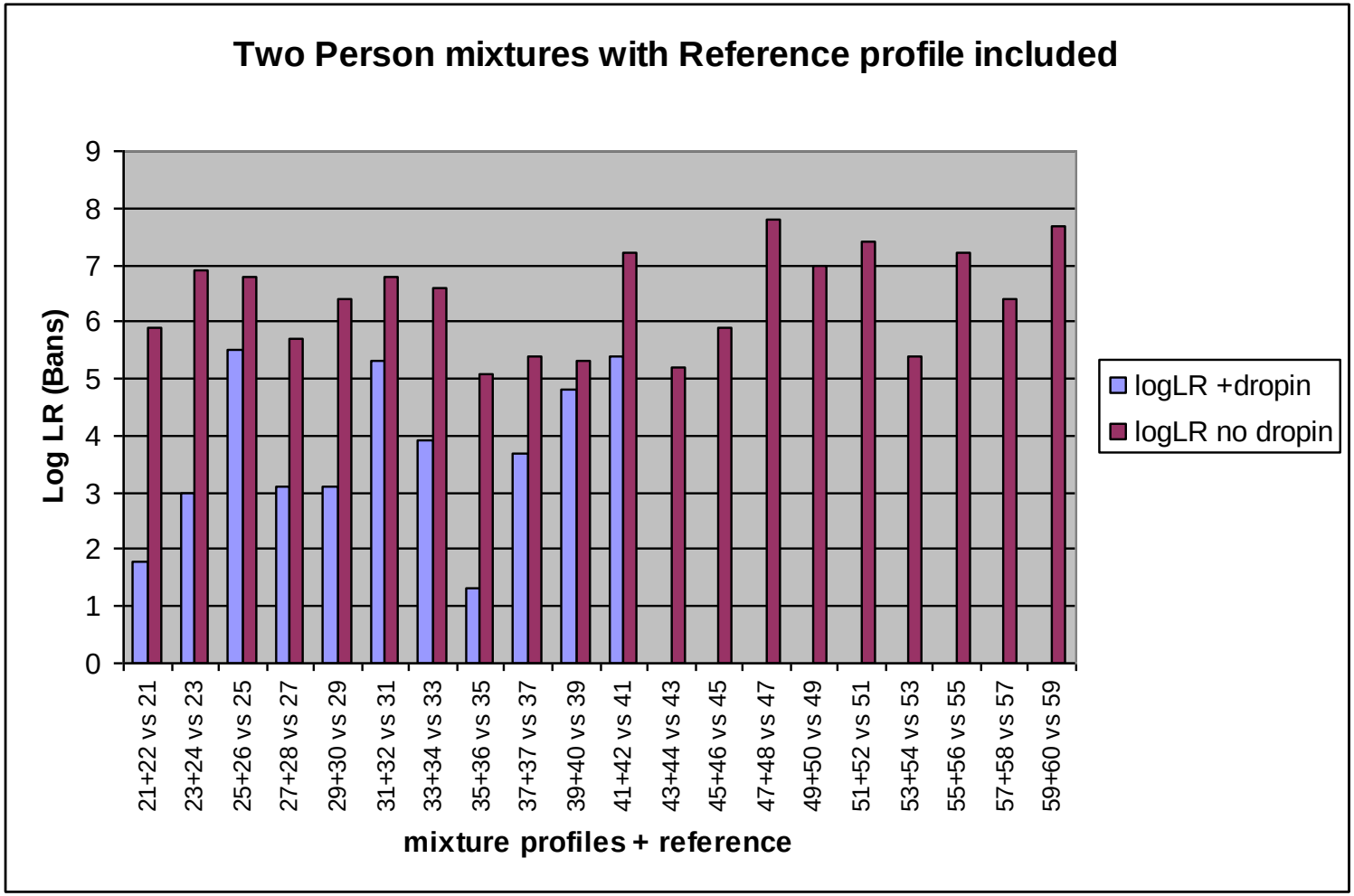


Effect of increasing numbers of replicates



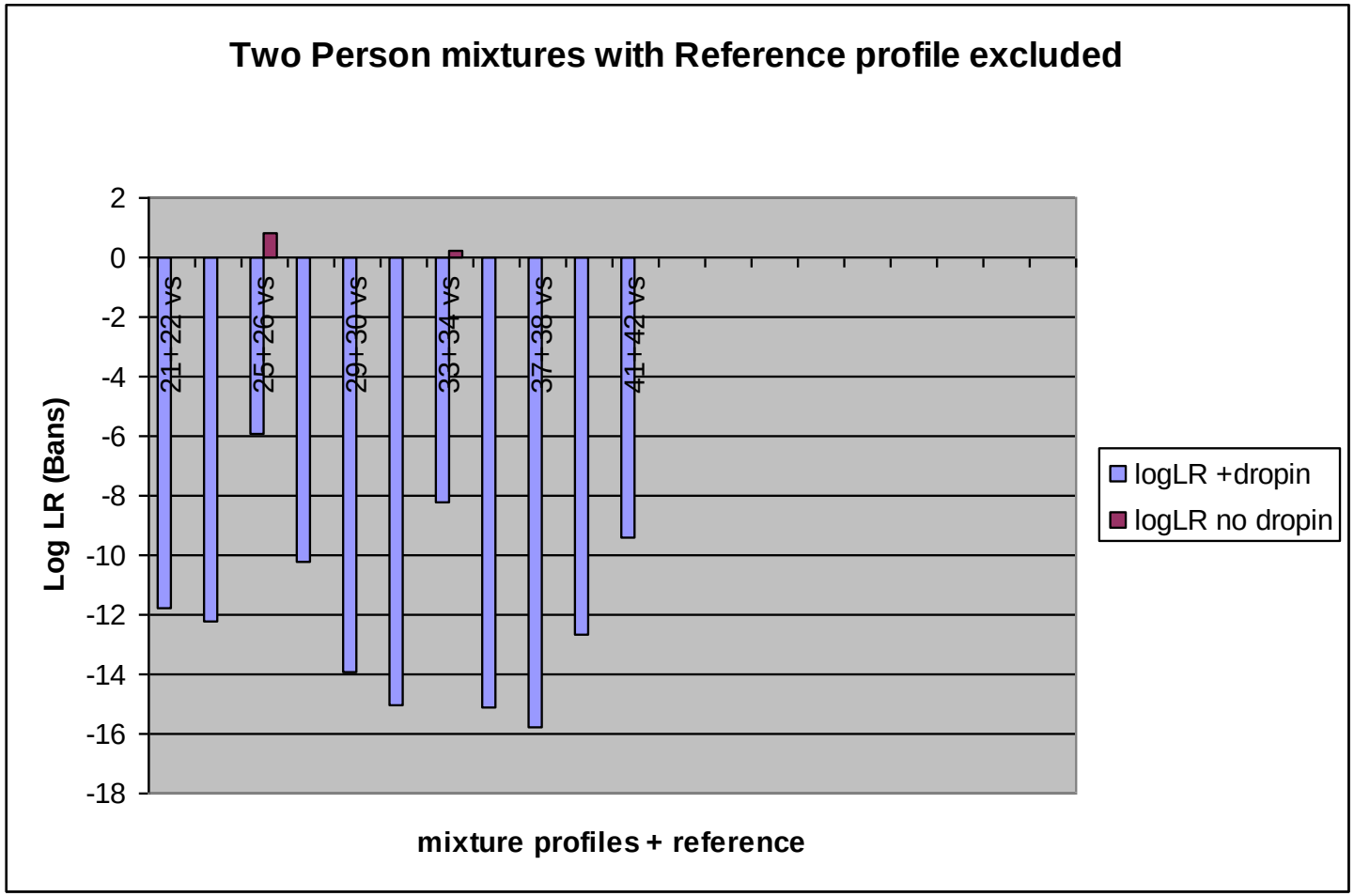
Cellmark Forensic Services

Two person mixtures with reference profile included



Cellmark Forensic Services

Two person mixtures with reference profile excluded



Cellmark Forensic Services

Advantages of likeLTD

- Provides an objective LR value in complex cases
- Can incorporate replicate PCR runs
- Freely available open source software
- Theoretical aspects published in peer reviewed journals
- Evidence successfully presented in court

Cellmark Forensic Services

Casework issues

- Limited to a maximum of 2 unknown contributors under Hp
- Memory requirements increase significantly with more loci
- Complex profiles can take several days to analyse on a standard desktop PC
- Can be difficult for non “R” code users

Cellmark Forensic Services

likeLTD publications

Interpreting low template DNA profiles

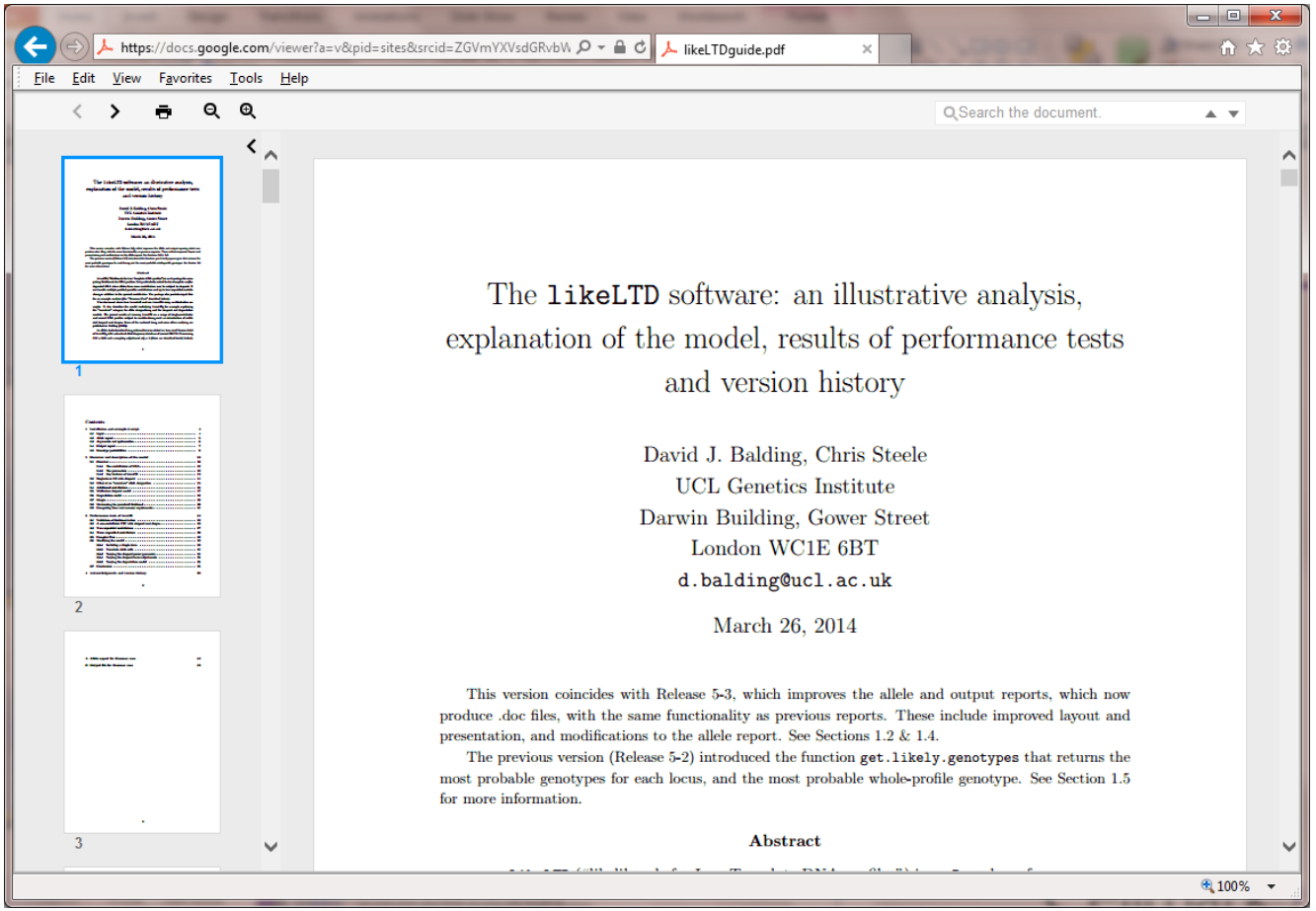
David J. Balding[a], John Buckleton[b]
[a] Department of Epidemiology and Public Health, Imperial College, St Mary's Campus, Norfolk Place, London W2 1PG, UK
[b] ESR Private Bag 92021, Auckland, New Zealand

Verifying likelihoods for low template DNA profiles using multiple replicates

Christopher D. Steele[a,*], Matthew Greenhalgh[b], David J. Balding[a]
[a] UCL Genetics Institute, Darwin Building, Gower Street, London WC1E 6BT, UK
[b] Orchid Cellmark Ltd., Abingdon Business Park, Blacklands Way, Abingdon OX14 1YX, UK

likeLTD guide

https://sites.google.com/site/baldingstatisticalgenetics/



Cellmark Forensic Services

likeLTD in court

- likeLTD evidence has been accepted without challenge in more than 10 trials in the UK
- There have been several admissibility challenges (voir dire) - all rejected
  - Evidence was originally ruled inadmissible in the case of R v MDS. At the subsequent retrial the evidence was again challenged but accepted

Cellmark Forensic Services

Summary

- likeLTD has been introduced into forensic casework following internal validation
- Provides objective LR values in complex mixture cases
- Evidence has been accepted in UK courts

Currently evaluating fully continuous probabilistic software

Cellmark Forensic Services

NIST DNA Analyst Webinar Series

Validation and use of likeLTD

Matthew Greenhalgh
Director of Forensic Science

18 September 2014